CRYSTAL STRUCTURE OF ESTROGEN RECEPTOR-β COMPLEX AND USES THEREOF

[0001] This application claims the benefit of U.S. Provisional Application No. 60/217,834 filed July 12, 2000.

Field of the Invention

[0002] The present invention relates to the crystal structure of the Estrogen Receptor- β (ER-β) complexed with genistein. This structure is critical for the design and selection of potent and selective agents which interact with ER-β, and particularly, the design of novel chemotherapeutic agents.

Background of the Invention

[0003] The beneficial effects of estrogen on bone maintenance, blood lipid profile, and the cardiovascular system are well known and account for the widespread use of hormone replacement therapy (HRT) in postmenopausal women (1). Estrogens and anti-estrogens affect several tissues, and the pattern of effects observed depends upon the particular ligand used (2). A major advance toward understanding the differential effects of various estrogenic compounds came with the recent discovery of an additional form of the estrogen receptor (3). The newly discovered receptor, named ER-β, is similar in sequence to the previously known form, now called ER-α. Mapping the distribution of ER-β and ER-α mRNA in normal and neoplastic tissues has provided an intriguing picture of differential expression patterns in different tissue types (4,5,6,7). The existence of clear-cut differences in receptor expression suggests that tissues could be targeted selectively with ligands selective for ER-α or ER-β.

[0004] Like all known nuclear receptors, estrogen receptors function as ligand-activated transcriptional factors and have a modular structure consisting of six discrete domains, named A-F. These domains mediate binding to DNA,

ligands and co-activators (8,9,10,11). The E domain of ER-α binds ligands such as 17β-estradiol and the phytoestrogen, genistein. The E-domains of ER-α and ER-β are 59% identical in sequence and have a predicted mass of approximately 25 kD. The natural ligand, 17β-estradiol, binds both with similar affinity. In contrast, genistein is selective, having 30 fold greater affinity for ER-β than for ER-α ((3) and H. Harris, unpublished observations).

[0005] The ligand binding domains (LBDs) of all studied nuclear receptors change conformation substantially upon ligand binding (12,13,14,15), particularly in the positioning of helix 12 (H12). In the case of ER-α, the position of H12 induced by the ligand depends on whether the ligand is an agonist (estradiol or diethylstilbestrol (DES)) or antagonist (raloxifene or tamoxifen). In the agonist complex, H12 packs against helices H3, H5, H6 and H11, forming a lid over the ligand. In this complex, H12 forms a wall perpendicular with and at one end of the co-activator binding groove formed by residues in H3, H4, H5 and the turn between H3 and H4. Peptides derived from the NR box II region (16,17,18,19) of the co-activator, GRIP1 can bind in this groove (11), suggesting this is an important aspect of transcriptional regulation. In contrast, steric hindrance from a bound antagonist displaces H12 so that it now binds in a hydrophobic groove formed by residues from helices 3 and 5. In this position, H12 binds to and occludes the co-activator recognition site, mimicking the interactions formed by the NR box II with the LBD and probably preventing modulation by co-activators. From these results it is clear that the structure of the bound ligand affects the overall structure of ER-α and its interactions with co-activators.

Summary of the Invention

[0006] The present invention provides a crystal of ER-β complexed with genistein, as well as the three dimensional structure of ER-β as derived by x-ray diffraction data of the ER-β/genistein crystal. Specifically, the three dimensional structure of ER-β is defined by the structural coordinates shown in

Figure 2, ± a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5Å. The structural coordinates of ER-β are useful for a number of applications, including, but not limited to, the visualization, identification and characterization of various active sites of ER-β, and the ER-β/genistein complex, including the genistein binding site. The active site structures may then be used to design various agents which interact with ER-β, as well as ER-β complexed with genistein or related molecules.

[0007] The present invention is also directed to an active site of a genistein binding protein or peptide, and preferably the genistein binding site of the ER-β, comprising the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å. Alternatively, the active site may include, in addition to the structural coordinates define above, the relative structural coordinates of amino acid residues VAL328, MET342, SER345, THR347, LYS348, LEU349, ALA350, ASP351, LEU354, MET357, TRP383, GLU385, VAL386, MET389, GLY390, LEU391, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, VAL535, TYR536 and LEU538 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å. The genistein active site may correspond to the configuration of ER-β in its state of association with an agent, preferably, genistein, or in its unbound state.

[0008] In another embodiment, the active site of a genistein binding protein or peptide, and preferably the genistein binding site of the ER-β, comprises the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, LEU391, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2

for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å. Alternatively, the active site may include, in addition to the structural coordinates define above, the relative structural coordinates of amino acid residues MET342, SER345, THR347, LYS348, ALA350, ASP351, MET357, TRP383, GLU385, VAL386, LEU387, MET389, GLY390, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, TYR536 and LEU538 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å. Here again, the genistein active site may correspond to the configuration of ER-β in its state of association with an agent, preferably, genistein, or in its unbound state.

[0009] In addition, the present invention provides a method for identifying an agent that interacts with ER-β, comprising the steps of: (a) generating a three dimensional model of ER-β using the relative structural coordinates according to Figure 2, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å; and (b) employing said three-dimensional model to design or select an agent that interacts with ER-β.

[0010] Still further the present invention provides a method for identifying an activator or inhibitor of a molecule or molecular complex comprising a genistein binding site, comprising the steps of: (a) generating a three dimensional model of said molecule or molecular complex comprising a genistein binding site using (i) the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, or (ii) the relative structural coordinates of amino acid residues MET343, LEU346, LEU349,

GLU353, MET384, LEU387, MET388, LEU391, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å; and (b) selecting or designing a candidate activator or inhibitor by performing computer fitting analysis of the candidate activator or inhibitor with the three dimensional model generated in step (a). In another embodiment, the relative structural coordinates according to (i) further comprises the relative structural coordinates of amino acid residues VAL328, MET342, SER345, THR347, LYS348, LEU349, ALA350, ASP351, LEU354, MET357, TRP383, GLU385, VAL386, MET389, GLY390, LEU391, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, VAL535, TYR536 and LEU538 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å. In yet another embodiment, the relative structural coordinates according to (ii) further comprises the relative structural coordinates of amino acid residues MET342, SER345, THR347, LYS348, ALA350, ASP351, MET357, TRP383, GLU385, VAL386, LEU387, MET389, GLY390, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, TYR536 and LEU538 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å.

[0011] Finally, the present invention provides agents, activators or inhibitors identified using the foregoing methods. Small molecules or other agents which inhibit or otherwise interfere with ER-β may be useful in the treatment of diseases associated with ER-β such as cancer.

[0012] Additional objects of the present invention will be apparent from the description which follows.

Brief Description of the Figures

[0013] Figure 1 provides a sequence alignment of ER-α with ER-β covering the ordered extent of ER-β. The numbering scheme used was chosen to be consistent with ER-α, such that the first ordered ER-β residue, 311 is residue 263 in the full length protein. Residues in helix 12 are underlined. The (*) symbols indicate the altered binding site residues.

[0014] Figure 2 provides the atomic structural coordinates for ER-β and genestein as derived by X-ray diffraction of an ER-β and genestein crystal complex. "Atom type" refers to the atom whose coordinates are being measured. "Residue" refers to the type of residue of which each measured atom is a part - i.e., amino acid, cofactor, ligand or solvent. The "x, y and z" coordinates indicate the Cartesian coordinates of each measured atom's location in the unit cell (Å). "Occ" indicates the occupancy factor. "B" indicates the "B-value", which is a measure of how mobile the atom is in the atomic structure (Å^2). Under "Residue type", "GEN C" refers to one molecule of genistein, "GEN D" refers to a second molecule of genistein, and "W" refers to water molecules.

Detailed Description of the Invention

[0015] As used herein, the following terms and phrases shall have the meanings set forth below:

[0016] Unless otherwise noted, Estrogen Receptor- β (ER-β) comprises the amino acid sequence depicted in Figure 1, including conservative substitutions.

[0017] "Genistein" is 4',5,7-trihydroxyisoflavone. A "genistein binding protein or peptide" is a protein or peptide that binds genistein and has a genistein binding site, and includes but is not limited to ER-β. A "molecule or molecular complex comprising a genistein binding site" includes ER-β and other molecules or molecular complexes having a genistein binding site.

[0018] Unless otherwise indicated, "protein" or "molecule" shall include a protein, protein domain, polypeptide or peptide.

[0019] "Structural coordinates" are the Cartesian coordinates corresponding to an atom's spatial relationship to other atoms in a molecule or molecular complex. Structural coordinates may be obtained using x-ray crystallography techniques or NMR techniques, or may be derived using molecular replacement analysis or homology modeling. Various software programs allow for the graphical representation of a set of structural coordinates to obtain a three dimensional representation of a molecule or molecular complex. The structural coordinates of the present invention may be modified from the original sets provided in Figure 2 by mathematical manipulation, such as by inversion or integer additions or subtractions. As such, it is recognized that the structural coordinates of the present invention are relative, and are in no way specifically limited by the actual x, y, z coordinates of Figure 2.

[0020] An "agent" shall include a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound or drug.

[0021] "Root mean square deviation" is the square root of the arithmetic mean of the squares of the deviations from the mean, and is a way of expressing deviation or variation from the structural coordinates described herein. The present invention includes all embodiments comprising conservative substitutions of the noted amino acid residues resulting in same structural coordinates within the stated root mean square deviation.

[0022] It will be obvious to the skilled practitioner that the numbering of the amino acid residues of ER-β may be different than that set forth herein, and may contain certain conservative amino acid substitutions that yield the same three dimensional structures as those defined by Figure 2 herein. Corresponding amino acids and conservative substitutions in other isoforms or analogues are easily identified by visual inspection of the relevant amino acid

sequences or by using commercially available homology software programs (e.g., MODELLAR, MSI, San Diego, CA).

[0023] "Conservative substitutions" are those amino acid substitutions which are functionally equivalent to the substituted amino acid residue, either by way of having similar polarity, steric arrangement, or by belonging to the same class as the substituted residue (e.g., hydrophobic, acidic or basic), and includes substitutions having an inconsequential effect on the three dimensional structure of ER-β with respect to the use of said structures for the identification and design of agents which interact with ER-β and genistein, as well as other proteins, peptides, molecules or molecular complexes comprising a genistein or ER-β binding site, for molecular replacement analyses and/or for homology modeling.

[0024] An "active site" refers to a region of a molecule or molecular complex that, as a result of its shape and charge potential, favorably interacts or associates with another agent (including, without limitation, a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound or drug) *via* various covalent and/or non-covalent binding forces. As such, an active site of the present invention may include, for example, the actual site of genistein binding with ER-β, as well as accessory binding sites adjacent or proximal to the actual site of genistein binding that nonetheless may affect ER-β activity upon interaction or association with a particular agent, either by direct interference with the actual site of genistein binding or by indirectly affecting the steric conformation or charge potential of the ER-β and thereby preventing or reducing binding of genistein to ER-β at the actual site of genistein binding. As used herein, an "active site" also includes analog residues of ER-β which exhibit observable NMR perturbations in the presence of a binding ligand, such as genistein. While such residues exhibiting observable NMR perturbations may not necessarily be in direct contact with or immediately proximate to ligand binding residues, they may be critical ER-β residues for rational drug design protocols.

[0025] The present invention first provides a crystallized complex comprising ER-β and genistein. In a particular embodiment, the amino acid sequence of ER-β is set forth in Figure 1, and includes conservative substitutions. The crystal complex of the present invention effectively diffracts X-rays for the determination of the structural coordinates of the complex of ER-β and genistein, and is characterized as having space group $P2_12_12_1$, and unit cell parameters of a=53.49Å, b=85.21Å, c=107.07Å. Further, the crystallized complex of the present invention consists of two molecules of ER-β each bound to a molecule of genistein.

[0026] Using a grown crystal complex of the present invention, X-ray diffraction data can be collected by a variety of means in order to obtain the atomic coordinates of the molecules in the crystallized complex. With the aid of specifically designed computer software, such crystallographic data can be used to generate a three dimensional structure of the molecules in the complex. Various methods used to generate and refine a three dimensional structure of a molecular structure are well known to those skilled in the art, and include, without limitation, multiwavelength anomalous dispersion (MAD), multiple isomorphous replacement, reciprocal space solvent flattening, molecular replacement, and single isomorphous replacement with anomalous scattering (SIRAS).

[0027] Accordingly, the present invention also provides the three dimensional structure of ER-β as derived by x-ray diffraction data of the ER-β/genistein crystal. Specifically, the three dimensional structure of ER-β is defined by the structural coordinates shown in Figure 2, ± a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. The structural coordinates of ER-β are useful for a number of applications, including, but not limited to, the visualization, identification and characterization of various active sites of ER-β, and the ER-β/genistein complex, including the genistein binding site. The active site structures may

then be used to design agents with interact with ER-β, as well as ER-β complexed with genistein or related molecules.

[0028] The present invention is also directed to an active site of a genistein binding protein or peptide, and preferably the genistein binding site of the ER-β, comprising the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. Alternatively, the active site may include, in addition to the structural coordinates define above, the relative structural coordinates of amino acid residues VAL328, MET342, SER345, THR347, LYS348, LEU349, ALA350, ASP351, LEU354, MET357, TRP383, GLU385, VAL386, MET389, GLY390, LEU391, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, VAL535, TYR536 and LEU538 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. The genistein active site may correspond to the configuration of ER-β in its state of association with an agent, preferably, genistein, or in its unbound state.

[0029] In another embodiment, the active site of a genistein binding protein or peptide, and preferably the genistein binding site of the ER-β, comprises the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, LEU391, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. Alternatively, the active site

may include, in addition to the structural coordinates define above, the relative structural coordinates of amino acid residues MET342, SER345, THR347, LYS348, ALA350, ASP351, MET357, TRP383, GLU385, VAL386, LEU387, MET389, GLY390, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, TYR536 and LEU538 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. Here again, the genistein active site may correspond to the configuration of ER-β in its state of association with an agent, preferably, genistein, or in its unbound state.

[0030] Another aspect of the present invention is directed to a method for identifying an agent that interacts with ER-β, comprising the steps of: (a) generating a three dimensional model of ER-β using the relative structural coordinates according to Figure 2, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å; and (b) employing said three-dimensional model to design or select an agent that interacts with ER-β. The agent may be identified using computer fitting analyses utilizing various computer software programs that evaluate the "fit" between the putative active site and the identified agent, by (a) generating a three dimensional model of the putative active site of a molecule or molecular complex using homology modeling or the atomic structural coordinates of the active site, and (b) determining the degree of association between the putative active site and the identified agent. Three dimensional models of the putative active site may be generated using any one of a number of methods known in the art, and include, but are not limited to, homology modeling as well as computer analysis of raw data generated using crystallographic or spectroscopy data. Computer programs used to generate such three dimensional models and/or perform the necessary

fitting analyses include, but are not limited to: GRID (Oxford University, Oxford, UK), MCSS (Molecular Simulations, San Diego, CA), AUTODOCK (Scripps Research Institute, La Jolla, CA), DOCK (University of California, San Francisco, CA), Flo99 (Thistlesoft, Morris Township, NJ), Ludi (Molecular Simulations, San Diego, CA), QUANTA (Molecular Simulations, San Diego, CA), Insight (Molecular Simulations, San Diego, CA), SYBYL (TRIPOS, Inc., St. Louis. MO) and LEAPFROG (TRIPOS, Inc., St. Louis, MO). The structural coordinates also may be used to visualize the three-dimensional structure of ER-β and the ER-β/genistein complex using MOLSCRIPT (28) and RASTER3D (29), for example.

[0031] The effect of such an agent identified by computer fitting analyses on ER-β activity may be further evaluated by contacting the identified agent with ER-β and measuring the effect of the agent on ER-β activity. Depending upon the action of the agent on the active site of ER-β, the agent may act either as an inhibitor or activator of ER-β activity. For example, enzymatic assays may be performed and the results analyzed to determine whether the agent is an inhibitor of ER-β and genistein (i.e., the agent may reduce or prevent binding affinity between ER-β and genistein) or an activator of ER-β and genistein (i.e., the agent may increase binding affinity between ER-β and genistein). Further tests may be performed to evaluate the potential therapeutic efficacy of the identified agent on conditions associated with ER-β such as cancer.

[0032] The present invention also provides a method for identifying an activator or inhibitor of a molecule or molecular complex comprising a genistein binding site, and preferably ER-β, comprising the steps of: (a) generating a three dimensional model of said molecule or molecular complex comprising a genistein binding site using (i) the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more

than 1.0Å, and most preferably not more than 0.5Å, or (ii) the relative structural coordinates of amino acid residues MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, LEU391, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å; and (b) selecting or designing a candidate activator or inhibitor by performing computer fitting analysis of the candidate activator or inhibitor with the three dimensional model generated in step (a). In another embodiment, the structural coordinates according to (i) further comprises the relative structural coordinates of amino acid residues VAL328, MET342, SER345, THR347, LYS348, LEU349, ALA350, ASP351, LEU354, MET357, TRP383, GLU385, VAL386, MET389, GLY390, LEU391, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, VAL535, TYR536 and LEU538 according to Figure 2 for monomer A of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. In yet another embodiment, the relative structural coordinates according to (ii) further comprises the relative structural coordinates of amino acid residues MET342, SER345, THR347, LYS348, ALA350, ASP351, MET357, TRP383, GLU385, VAL386, LEU387, MET389, GLY390, MET392, LEU402, ILE403, ALA405, LEU408, VAL418, GLU419, GLY420, LEU422, GLU423, PHE425, LEU428, ALA516, SER517, LYS519, MET521, GLU522, LEU525, ASN526, MET527, LYS528, VAL533, TYR536 and LEU538 according to Figure 2 for monomer B of ER-β, ± a root mean square deviation from the backbone atoms of said amino acids of not more than 1.5Å, preferably not more than 1.0Å, and most preferably not more than 0.5Å. Once the candidate activator or inhibitor is obtained or synthesized, the candidate activator or inhibitor may be contacted

with the molecule or molecular complex, and the effect the candidate activator or inhibitor has on said molecule or molecular complex may be determined. Preferably, the candidate activator or inhibitor is contacted with the molecule or molecule complex in the presence of genistein (or a molecule or a molecular complex comprising genistein) in order to determine the effect the candidate activator or inhibitor has on binding of the molecule or molecular complex to genistein.

[0033] Various molecular analysis and rational drug design techniques are further disclosed in U.S. Patent Nos. 5,834,228, 5,939,528 and 5,865,116, as well as in PCT Application No. PCT/US98/16879, published WO 99/09148, the contents of which are hereby incorporated by reference.

[0034] The present invention is also directed to the agents, activators or inhibitors identified using the foregoing methods. Such agents, activators or inhibitors may be a protein, polypeptide, peptide, nucleic acid, including DNA or RNA, molecule, compound, or drug. Small molecules or other agents which inhibit or otherwise interfere with ER-β and genistein may be useful in the treatment of diseases associated with ER-β such as cancer.

[0035] In addition, the present invention is directed to a method for determining the three dimensional structure of a molecule or molecular complex whose structure is unknown, comprising the steps of obtaining crystals of the molecule or molecular complex whose structure is unknown and generating X-ray diffraction data from the crystallized molecule or molecular complex. The X-ray diffraction data from the molecule or molecular complex is then compared with the known three dimensional structure determined from any of the aforementioned crystals of the present invention. Then, the known three dimensional structure determined from the crystals of the present invention is "conformed" using molecular replacement analysis to the X-ray diffraction data from the crystallized molecule or molecular complex. Alternatively, spectroscopic data or homology modeling may be used to generate a putative three dimensional structure for the molecule or molecular complex, and the

putative structure is refined by conformation to the known three dimensional structure determined from any of the crystals of the present invention.

[0036] The present invention may be better understood by reference to the following non-limiting Example. The following Example is presented in order to more fully illustrate the preferred embodiments of the invention, and should in no way be construed as limiting the scope of the present invention.

Example 1

[0037] We describe the 1.8Å crystal structure of the recently discovered nuclear hormone receptor, ER-β, in complex with genistein, an agonistic phytoestrogen. The overall structure is similar to that of previously described ER-α complexes, with genistein occupying a central cavity similar to that of ER-α. Minor differences between the two cavities account for genistein's 30 fold selectivity for ER-β over ER-α. Surprisingly, helix 12 in the complex of ER-β with genistein (an agonist) runs in the same direction, although in a different position, as helix 12 of ER-α bound to raloxifene (an antagonist). This suggests different mechanisms of agonism/antagonism for ER-α and ER-β.

1. Methods and Methods

[0038] *Cloning, Expression and Purification.* Human ER-β cDNA (21) was generated from human testis RNA by RT-PCR and cloned into the mammalian expression vector pcDNA3. The LBD of human ER-β was then PCR amplified from the cloned cDNA and inserted into the *E. coli* expression vector pET16b between the NcoI and XhoI restriction sites. The expressed LBD thus has the following sequence: MD[D_{261}-L_{500}]DD.

[0039] Frozen cells were lysed by two cycles in a French press (SLM Instruments). The protein was loaded on an estradiol-Sepharose column and the column was washed with 100 mL of 10 mM Tris-HCl, pH 7.5 containing 0.5 M NaCl, 5 mM dithiothreitol, and 1mM EDTA. The column was then reequilibrated with 10 mM Tris-HCl, pH 7.5, 0.2M NaCl, 1 mM EDTA (buffer

A), and accessible cysteines modified by 5 mM iodoacetic acid in the same buffer. The protein was eluted with 200 μM genistein and 5 mM DTT then passed over a G3000 SW TosoHaas size exclusion column equilibrated with buffer A. Mass spectroscopy (MALDI) showed that two cysteine residues had been modified by carboxymethylation.

[0040] *Crystallization and Data Collection.* The ERβ/genistein complex was concentrated to 12 mg/mL in 0.2M NaCl, 1mM EDTA, 5mM DTT, 10mM Tris-HCl pH 7.5 buffer. Crystals were grown using vapor diffusion at 4°C over wells containing 12% PEG2000 mono-methyl ether buffered with 0.1M MES (2-[N-morpholino] ethane sulphonic acid) at pH 6.0. Crystals were cooled to 100K in 8.5% PEG 2000, 3.6% PEG 8000, 5% glycerol, 0.13M MES pH 6.0, 0.02M sodium cacodylate pH 6.5, and 0.04M calcium acetate. The space group is $P2_12_12_1$ with cell parameters a=53.49Å, b=85.21Å, c=107.07Å. Diffraction data were collected on station 5.0.2 at the Advanced Light Source, Berkeley, using a Quantum-4 CCD detector (Area Detector Systems), then reduced using DENZO/SCALEPACK (22), giving the statistics in Table 1.

[0041] *Phasing and Refinement.* The structure was solved using AMORE (23) molecular replacement with ER-αLBD (24) (without bound ligand, the loop H8-H9, C and N terminal helices) as a search model. The resulting 2Fo-Fc map showed clear density for the bound genistein not included in the phasing model.

[0042] BUSTER (25)/TNT (26) was used to generate maximum entropy omit maps to reduce model bias and generate a more detailed map for the bound ligand. REFMAC (27) was used for all further refinement of the model, giving the statistics in Table 1. The model consists of residues Leu311-Ala549, with the first ten residues, loop Tyr459-Ala468, and the last three residues in disordered regions. The dimer has two bound genistein molecules and 189

ordered water molecules. Electron density for the cysteine modifications was poor and was therefore not modeled.

2. Results

[0043] Like ER-α, ER-β has a predominantly globular structure formed by anti-parallel α helices arranged in three layers, and a short two-stranded β ribbon. A cavity is formed in the core of the protein which becomes occupied by ligand. In the structure reported here, the cavity is occupied by a molecule of genistein, which is completely buried and forms hydrogen bonds with a single buried water molecule. The protein forms a non-crystallographic dimer with a large interface formed by helices H9, H10 and H11 (not shown), consistent with size exclusion chromatography studies on the protein.

[0044] The overall structure of the ER-β-genistein complex is very similar to previously reported ER-α structures (10) (not shown). Superimposing 426 of 461 C-α coordinates of the dimer, the RMS differences with ER-α are only 0.53Å and 0.57Å for complexes with 17β-estradiol and raloxifene, respectively. There are, however, significant differences that may account for the selectivity of ER-β over ER-α when binding certain ligands and co-activator peptides (20). The most striking difference is in the position of H12. The position of this helix in ER-α has been found to vary depending upon the ligand bound. In the agonist (17β-estradiol) bound structure (not shown), H12 lies over the ligand and encapsulates it within the core of the protein. This conformation facilitates binding of co-activator peptides in a hydrophobic groove just below H12 (not shown), formed by residues from H3, H4 and H5. When ER-α binds an antagonist such as raloxifene, part of the ligand prevents H12 from occupying this location (not shown). Instead, H12 occupies the co-activator hydrophobic groove, thus preventing co-activation. The location of H12 in complexed ER-β is different from that of either ER-α-agonist or ER-α-antagonist complexes (not shown). H12 in ER-β runs in the same direction as H12 in ER-α-antagonist but does not cover the co-activator binding site. It occupies roughly the same space

as H12 of the ER-α-agonist and buries the bound ligand. This unexpected position of H12 was observed in crystals of space group $P2_12_12_1$ (described here) and $P3_221$ (2.4Å, refined to an R-value of 0.238 and a free R-value of 0.286, not described here). The equivalent position of H12 observed in two crystal forms suggests that this location is correct.

[0045] Genistein is bound in the hydrophobic core of the protein and is completely shielded from the bulk solvent by H3, H6, strand S1, H7, H8, H11 and H12. H12 appears to form a lid over the filled cavity. Bound genistein superimposes well with 17β-estradiol bound to ER-α (not shown) such that the phenolic moieties are in similar positions and the fused rings lie over the puckered C/D rings of the steroid. The position of the 7OH hydroxyl group of genistein in the same as that of 17-OH of 17β-estradiol, allowing formation of similar hydrogen bonds. The phenolic hydroxyl (4'OH) of genistein hydrogen bonds with the OE2 of Glu353(2.63Å), the NH2 of Arg394 (2.93Å) and a highly ordered water molecule (3.05Å). This ordered water molecule was also found in ER-α when bound to agonist or antagonist and must therefore be considered part of the binding site. The last hydrogen bond genistein forms is between its 7-OH and the ND1 of His523 (2.62Å). The position of His523 is stabilized by an interaction with the backbone carbonyl of Gly419 at the N-terminal end of H8.

[0046] There are very few differences between the binding cavities that can account for genistein's 30 fold preference for binding to ER-β. The substitution of Met421 to Ile in ER-β is likely the most significant. In the complex of ER-α with 17β-estradiol, the Met Sd lies 4.4Å from the C16 atom in the puckered "D" ring. Upon superimposition of genistein into the ER-α binding cavity, the Met Sd would now appear to lie only 3.9Å from genistein's 5-OH group, which is unable to move away due to the planarity of its ring. This close interaction is unfavorable not only sterically, but also electrostatically, due to close proximity of the small negative charges on the Sd and the 5-OH group. A superimposition of genistein against the ER-α/DES complex demonstrates a similar (and slightly more severe) steric and electrostatic clash. In contrast, ER-

β places the Cg1 and Cd1 of Ile421 4.2Å away from atoms in the ring thereby promoting more favorable van der Waals interactions. The other difference between the cavities that could affect selectivity is Leu384 to Met. The Met384 Ce of ER-β extends far into the binding cavity and provides favorable interactions with genistein, while placing the Met Sd of ER-β in the same position as the Leu Cd1 of ER-α. When looking down on the fused "B" ring from Met384, the Sd appears to be directly above the "B" ring at a distance of 4.2Å, making favorable interactions. The Sd-Ce bond projects in the same plane as the B-C fused ring and places the terminal carbon 3.8Å from the genistein O1 atom, making a good van der Waals interaction that would not be present in ER-α. Packing of Met Sd and Ce against the face of aromatic residues is common in protein structures and must therefore be considered as a stabilising interaction. The final difference between the binding cavities relates to the position of H12 and probably does not significantly impact the selectivity. In both forms of ER, leucines in H12 project down towards the bound ligand and seal the binding site. In ER-β Leu538 plays this role, whereas in ER-α it is Leu540. The difference is due to different direction of H12 in the two structures.

[0047] The position of H12 in ER-α has been used as the structural hallmark of a bound agonist versus antagonist (not shown). In ER-β, H12 runs in the opposite direction as in agonist ER-α H12. However, it has several non-equivalent, but superimposable side chains. Specifically, ER-α Met543 is replaced by ER-β Leu539 and ER-α Leu539 with ER-β Met542. Although the position of the ER-β helix can be roughly superimposed on ER-α H12, it does partially occlude the co-activator binding site. Co-activators may bind ER-β in a slightly different manner than ER-α, or ER-β's H12 may move upon binding of co-activators, to allow full access to the binding groove.

[0048] Why does H12 occupy a different location in agonist-bound ER-α versus ER-β? The only amino acid difference is Asp545 to Asn. This difference seems unlikely to explain the different positions of H12, as either residue is exposed to solvent and appears not to affect the positioning of H12. The key

may lie in the residues just upstream of the helix which also move upon agonist/ antagonist binding. Leu536 of ER-α, which makes favorable interactions in a hydrophobic cavity is replaced by Val in ER-β. Val cannot make comparable contacts. Two other differences, Gly344 to Met and Asn348 to Lys, influence the nature of the H12 binding surface in this same region and may affect the position of the loop just upstream of H12. Changes in the position of H12 could explain the selective binding of co-activators to ER-α and ER-β. Steroid receptor coactivator-3 (SRC-3) binds approximately 700 fold tighter to ER-α whereas SRC-1 preferentially activates ER-β (20).

[0049] The structure of ER-β in complex with genistein has enhanced our understanding of mechanisms of estrogen receptor agonism and co-activation. The structure has revealed the basis for the 30 fold selectivity of genistein for ER-β over ER-α and helps explain differences in co-activator binding. These insights should help in the design of more selective and therapeutically useful agonists or antagonists.

Table 1. Data Collection and Refinement Statistics.

Resolution limits (Å)	15.0-1.8(1.83-1.8)
Rmerge (%)[2]	5.7(23.0)
Unique reflections	45530(1977)
Total observations	287034
Completeness (%)	98.8(88.1)
<I/sigma>	26.2(4.1)
Refinement	
Refinement reflections	43176
R-value	22.4
Free reflections[3]	2294
Rfree	26.4
Average B – factor(Å2)[4]	31.5
R.m.s. differences[5]	
Main chain(Å2)	1.94
Side Chain(Å2)	3.29
R.m.s. deviations	
Bond lengths (Å)	0.014
Angles distance (Å)	0.033
Ramachandran Plot	
Most favored regions(%)	96.7

1. Values in parentheses refer to the highest resolution shell.
2. Rmerge = S | I - <I> |/S<I> where I is the observed intensity and <I> is the average of the symmetry mates.
3. 5% of the data were randomly selected, not used in refinement and used for the calculation of the Free R-value.
4. All atoms in the structure.
5. R.m.s. differences in B factors between bonded atoms.

Table 2*

Residues of ER-β That Interact With Genistein at 0-4 Å

From First Monomer of ER-β
MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524

From Second Monomer of ER-β
MET343, LEU346, LEU349, GLU353, MET384, LEU387, MET388, LEU391, ARG394, PHE404, ILE421, ILE424, GLY520, HIS523 and LEU524

Residues of ER-β That Interact With Genistein at 4-8 Å

From First Monomer of ER-β
VAL328, MET342, MET343, SER345, LEU346, THR347, LYS348, LEU349, ALA350, ASP351, GLU353, LEU354, MET357, TRP383, MET384, GLU385, VAL386, LEU387, MET388, MET389, GLY390, LEU391, MET392, ARG394, LEU402, ILE403, PHE404, ALA405, LEU408, VAL418, GLU419, GLY420, ILE421, LEU422, GLU423, ILE424, PHE425, LEU428, ALA516, SER517, LYS519, GLY520, MET521, GLU522, HIS523, LEU524, LEU525, ASN526, MET527, LYS528, VAL533, VAL535, TYR536 and LEU538

From Second Monomer of ER-β
MET342, MET343, SER345, LEU346, THR347, LYS348, LEU349, ALA350, ASP351, GLU353, MET357, TRP383, MET384, GLU385, VAL386, LEU387, MET388, MET389, GLY390, LEU391, MET392, ARG394, LEU402, ILE403, PHE404, ALA405, LEU408, VAL418, GLU419, GLY420, ILE421, LEU422, GLU423, ILE424, PHE425, LEU428, ALA516, SER517, LYS519, GLY520, MET521, GLU522, HIS523, LEU524, LEU525, ASN526, MET527, LYS528, VAL533, TYR536 and LEU538

* The ER-β/genistein complex molecular structure is a dimer with each monomer of ER-β binding to one genistein molecule.

References

[0050] 1. Jenning, T.S., Creasman, W.T., Estrogens and Antiestrogens: Basic and Clinical Aspects; Lindsay, R.; Dempster, D.W., Jordan, V.C., Eds; Lippencott-Raven: Philadelphia, 1997, pp. 223.

[0051] 2. Glasebrook, A.L., *et al.*, Proc Natl Acad Sci (USA) 94:14105-14110 (1997).

[0052] 3. Kuiper GGJM, *et al.* Proc Natl Acad Sci (USA) 93:5925-5930 (1996).

[0053] 4. Couse JF, *et al.*, Endocrinology 138(11):4613-4621 (1997).

[0054] 5. Shughrue PJ, *et al.*, Steroids 61:678-681 (1996).

[0055] 6. Lau K, *et al.*, Endocrinology 139(1):424-427 (1998).

[0056] 7. Enmark E, *et al.*, J Clin Endocrinol Metab 82(12):4258-4265 (1997).

[0057] 8. Tsai, M.J. and O'Malley, B.W., Annu. Rev. Biochem 63:451-486 (1994).

[0058] 9. Evans, R. M., Science 240: 889-895 (1988).

[0059] 10. Horwitz, K.B., *et al.*, Mol. Endocrinol. 10: 1167-1177 (1996).

[0060] 11. Glass, C.K., *et al.*, Curr. Opin. Cell Biol. 9: 222-232 (1997).

[0061] 12. Brzozowski, A., *et al.*, Nature 389: 753-758 (1997).

[0062] 13. Shiau, A., *et al.*, Cell. 95: 927-937 (1998).

[0063] 14. Bourguet, W., *et al.*, Nature 375: 377-382 (1995).

[0064] 15. Renaud, J., *et al.*, Nature 378: 681-689 (1995).

[0065] 16. Le Dourin, B. *et al.*, EMBO J15: 6701-6715 (1996).

[0066] 17. Heery, D., *et al.*, Nature 387: 733-736 (1997).

[0067] 18. Torchia, J. *et al.*, Nature 387: 677-684 (1997).

[0068] 19. Ding, S., *et al.*, Mol. Endocrinol. 12: 302-313 (1998).

[0069] 20. Suen, C-S., *et al.*, J. Biol. Chem. 273: 27645-27653 (1998).

[0070] 21. Mosselman, S., *et al.*, FEBS letters 392: 49-53 (1996).

[0071] 22. Otwinowski, Z. & Minor, W., Methods Enzymol., 276: 307-326 (1997).

[0072] 23. Collaborative Computational Project, Number 4. (1994). Acta Cryst. D50.

[0073] 24. D.M. Tanenbaum, *et al.*, PNAS 95: 5998-6003 (1998).

[0074] 25. Bricogne, G., Acta Cryst. D49: 37-60 (1993).

[0075] 26. Tronrud, D.E., (1997) Methods in Enzymology. 277B eds. Charlie Carter and Robert Sweet.

[0076] 27. Murshudov, G.N., *et al.*, Acta. Cryst. D53: 240-255 (1997).

[0077] 28. Kraulis, PJ, J. Appl. Crystallogr. 24: 946-950 (1991).

[0078] 29. Bacon, D.J. and Anderson, W.F., J. Mol. Graph. 6: 219-220 (1998).

[0079] All publications mentioned herein above, whether to issued patents, pending applications, published articles, or otherwise, are hereby incorporated by reference in their entirety. While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of the disclosure that various changes in form and detail can be made without departing from the true scope of the invention in the appended claims.